6
CAUSALITY AND CAUSAL INFERENCE IN MEDICINE

Julian Reiss

1. Inferring and Interpreting Causal Claims

Causal claims such as (1) "HPV types 16 and 18 are responsible for 70% of cervical cancers," (2) "Antipsychotic medications are effective in treating acute psychosis and reducing the risk of future psychotic episodes," (3) "Keeping weight under control, exercising, eating healthily, and not smoking prevent diabetes type-2," or (4) "Hal's exposure to vinyl chloride monomer caused his angiosarcoma" are widely used in medicine, epidemiology, public health, and elsewhere. Arguably, this is for good reason: causal claims allow us to *explain* medical outcomes, both at the population level (claim 1) and the individual level (claim 4); they can help to *predict* outcomes (claim 2); they *underwrite treatment decisions* (claim 2) and *public health policies* (claim 3); and they can help to *attribute responsibility* (join claim 4 to the claim that Hal's employer behaved negligently by exposing him)—all important aims of the health sciences and beyond.

Knowledge of causal claims is a good thing, then. However, their usefulness depends on the extent to which we can reliably learn causal relationships from data, the extent to which what we can learn from the data is unambiguous, and the extent to which causal knowledge actually does help to realize these more ultimate purposes. This chapter will examine these three issues in turn, starting with the learning of causal relationships from data, then moving on to the interpretation of causal claims, and finally addressing the usefulness issue. Throughout, I will use vinyl chloride carcinogenicity as a central example. Vinyl chloride is a compound widely used in the production of polyvinyl chloride (PVC). Before it was classified as a human carcinogen by the International Agency for Research on Cancer (IARC), the U.S. Environmental Protection Agency (EPA), and similar bodies in the 1970s, and then regulations were tightened in response, workers in PVC manufacturing plants were often exposed to the substance over long periods of time and suffered adverse health consequences, including the very rare condition angiosarcoma of the liver.

2. Inferring Causal Claims

It would be as imprudent as it would be uncouth to talk about causal inference in medicine without beginning the discussion with the so-called Hill criteria for causation, named after Sir Austin Bradford Hill (1897–1991), a British epidemiologist and statistician. He was concerned with how to distinguish genuinely causal from spurious associations. Medical researchers at the time were fully aware that an association between a risk factor such as diet and a medical outcome such as diabetes could be due to numerous reasons other than the risk factor causing

the disease. Hill proposed the following criteria for causation (Hill 1965; the descriptions are paraphrased rather than directly quoted):

(1) *Strength*. The stronger the association, the more likely it is causal.
(2) *Consistency*. If the association has repeatedly been observed by different persons, in different places, circumstances, and times, it is more likely to be causal.
(3) *Specificity*. The more specifically the risk factor, outcome, or both can be defined, the more likely an association between them is causal.
(4) *Temporality*. The more closely the temporal dimension of an association is aligned with our expectations, the more likely it is causal.
(5) *Biological gradient*. If there is a dose-response curve, the association is more likely to be causal.
(6) *Plausibility*. If the causal relation appears plausible on the basis of our background knowledge at the time, it is more likely to be causal.
(7) *Coherence*. The causal interpretation of the association should be coherent with general known facts about the natural history of the disease and its biology.
(8) *Experiment*. If the association has been established in an experiment or quasi-experiment, it is more likely to be causal.
(9) *Analogy*. If there is a strong analogy between the risk factor and a known cause of the disease, it is more likely to be causal.

None of these "viewpoints," as Hill sometimes calls them, is either necessary or sufficient for causality, nor is the conjunction of all nine sufficient. A strong association can be due to a confounder and a weak one causal. Plausibility is, as duly noted by Hill, relative to the biological knowledge of the day, which may be imperfect. There may not exist any analogies or experimental evidence. Nevertheless, each "viewpoint" can be regarded as a fallible indicator—rather than a strict criterion—of causality.

To see how the items on Hill's list can perform the role of fallible indicators of causality, it is useful to distinguish two competing approaches to causal reasoning in the biomedical sciences: the experimentalist on the one hand and the inferentialist on the other (cf. Parascandola 2004). The experimentalist approach maintains that randomized experiments are the "gold standard" of causal inference and discounts evidence about causal claims from other sources. Evidence-based medicine and other movements that carry the "evidence-based" label are rooted in the experimentalist approach. Inferentialism, by contrast, holds that causal claims are inferred from diverse bodies of evidence—bioassays, laboratory experiments with animal models, cohort and case-controlled studies, case reports, clinical trials—using pragmatic guidelines such as Hill's. The approach is very widely used in biomedical research, but inferentialists tend to be less vocal than their evidence-based colleagues.

One way to defend experimentalism is to assume a specific interpretation of causality and then proceed to show that under that interpretation of causality, positive results of certain kinds of experiments guarantee the truth of the associated causal claim. A view of causality that has been very popular recently is James Woodward's, according to which one variable X (directly) causes another variable Y if and only if there is a possible intervention (intervention variable I) on X that changes Y or its likelihood of occurring (Woodward 2003: 55). Intervention variable I, in turn, has the following characteristics (Woodward 2003: 98):

I1. I causes X.
I2. I acts as a switch for all the other variables that cause X. That is, certain values of I are such that when I attains those values, X ceases to depend on the values of other variables that cause X and instead depends only on the value taken by I.

I3. Any directed path from *I* to *Y* goes through *X*. That is, *I* does not directly cause *Y* and is not a cause of any causes of *Y* that are distinct from *X* except, of course, for those causes of *Y*, if any, that are built into the *I-X-Y* connection itself; that is, except for (a) any causes of *Y* that are effects of *X* (i.e., variables that are causally between *X* and *Y*) and (b) any causes of *Y* that are between *I* and *X* and have no effect on *Y* independently of *X*.
I4. *I* is (statistically) independent of any variable *Z* that causes *Y* and that is on a directed path that does not go through *X*.

In a biomedical experiment, the intervention is the assignment of a member of a test population (e.g., a population of animal models) to a treatment or control group. In one experiment, 360 Swiss mice were exposed to different concentrations of vinyl chloride (VC) 4 hours daily on 5 days per week for 30 weeks (Maltoni and Lefemine 1974). The assignment to a treatment group causes the level of exposure (I1). The level of exposure to VC is controlled by the experiment so that it no longer depends on other variables (e.g., proximity to a PVC manufacturing plant; I2). The assignment to a treatment group does not cause cancer through a mechanism that bypasses exposure to VC (I3). This condition would be violated, for example, if different treatment groups received different diets, which in turn affected cancer rates. Finally, the assignment to a treatment group is statistically independent of other causes of cancer (I4). This condition would be violated, for example, if different strains of mice, which have different degrees of cancer susceptibility, were used in different treatment groups. The four conditions are illustrated in Figure 6.1.

In clinical trials on human subjects, randomization is an intervention in Woodward's sense. If *X* is treatment status (with values x_t = test treatment and x_c = control), and *Y* a variable measuring the difference in medical outcome between treatment and control group, then randomization causes *X*, treatment status, to assume its value x_t or x_c: the outcome of the randomization process determines whether a patient will be in the treatment or the control group (I1). Clinical control will ensure that only patients who are in the treatment group will receive the treatment (I2). The outcome of the randomization process will not have a direct effect on the value of *Y* because *Y* is defined as the *difference* in medical outcome between the two groups and all participants of the trial—patients, doctors, nurses, analysts—are blinded with respect to treatment status (I3). Finally, successful randomization will guarantee (at least for large samples) statistical independence from other variables responsible for the medical outcome (I4). Thus, under a Woodwardian conception of causality, if in a randomized experiment that fulfills criteria (I1)–(I4) *X* and *Y* are associated, then it must be the case that *X* causes *Y*. This would appear to support experimentalism.

Critics of experimentalism say two things in response. First, few real clinical trials strictly fulfill criteria (I1)–(I4). Compliance (whether or not a patient takes the assigned treatment) is rarely perfect, and occasionally treatments are shared between the two groups. Withholding treatment status from participants is often not possible or, if it is possible, treatment status

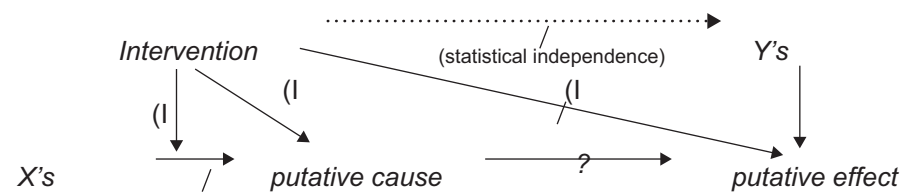

Figure 6.1 Woodward's characterization of an intervention variable

might be revealed if the treatment is actually effective. Randomization guarantees statistical independence of other causes of the medical outcome only in infinitely large samples, and test populations are often small and never infinite. The critics charge that while experimentalists have a good answer why ideal randomized experiments guarantee true causal conclusions, they remain silent about the conditions under which reliable causal judgments can be made under realistic conditions when experiments cannot be implemented ideally or when the causal claim is not open to experimental test. Second, they point out that the bulk of causal knowledge in medicine has been established non-experimentally. As one article published in the *British Medical Journal* put it ironically, we do not need a randomized controlled trial "to determine whether parachutes are effective in preventing major trauma related to gravitational challenge" (Smith and Pell 2003). Critics maintain that experimentalism makes it mysterious that other sources of evidence such as observational studies should reliably support causal claims.

Inferentialists regard causal inference as analogous to medical diagnosis. The presence of most diseases cannot be directly observed but only inferred on the basis of symptoms. Causality can analogously be inferred from "symptoms," albeit, in general, on the population level rather than the individual level. An association between a risk factor and a medical outcome can thus be regarded as a symptom of the existence of a causal relation between factor and outcome. Medical symptoms rarely point unequivocally to a single disease. Likewise, the symptoms of causality have alternative explanations such as confounding (the production of an association by a variable associated with risk factor and medical outcome), selection bias (the determination of treatment status by the patient), experimenter bias (the determination of treatment status by the experimenter), attrition bias (patients' premature exit from the trial at different rates between treatment and control arm), diagnostic error/mismeasurement, and so on. Thus, like a medical diagnostician, the causal researcher comes to a judgment about the hypothesis at stake only after looking at a diverse body of evidence, one function of which is to rule out such alternative explanations.

Just as there is no "golden symptom"—a type of test that reliably indicates the presence of a disease no matter what the disease is—there is no gold standard of evidence according to the inferentialist. There are only bodies of evidence that, if the parts fit together in the right way, can make a convincing case for or against a causal claim.

Hill's viewpoints can play the role of pragmatic criteria that help the inferentialist to come to a judgment concerning the causal claim. For example, a strong association [Hill's no. (1)] certainly does not prove a causal relationship—a confounded relationship may well be strong. (To use a philosopher's favorite example, drops in barometer readings are strongly associated with the occurrence of storms; the relationship is nevertheless confounded by atmospheric pressure.) However, if it is known that the most likely confounder cannot (or most likely does not) produce an association of that size, this alternative can be ruled out. This consideration helped to eliminate R. A. Fisher's "constitutional hypothesis" in the 1950s. According to this hypothesis, a single genetic factor predisposed people to both lung cancer and taking up smoking, thus accounting for the association between smoking and lung cancer without smoking necessarily being a cause of lung cancer. Although genetic factors were known to play a role in cancer susceptibility, they could not explain the 60-fold increased risk observed in the data.

On the other hand, weak associations do not disprove a causal relationship either. Vinyl chloride was classified as carcinogenic by the IARC in 1974 (IARC 1974a), but despite a surge in industrial use of the compound, only very few additional cases of cancer were observed. Carcinogenicity could nevertheless be established because exposed individuals developed angiosarcomas of the liver. These are so rare that chance or other confounders can hardly account for the coincidence of this specific condition [Hill's no. (3)] among exposed individuals.

Experimentation [Hill's no. (8)] plays a role in the inferentialist approach, but like the other items on Hill's list, it is neither necessary nor sufficient for establishing causality. Clinical trials cannot be used to establish carcinogenicity for ethical reasons. Vinyl chloride had been established experimentally to cause cancer in animal models before epidemiological studies confirmed its carcinogenicity in humans (IARC 2014), but even though all human carcinogens have some animal models, there is no guarantee that a substance that causes cancer in animals is also harmful to humans.

Although a well-designed randomized trial eliminates a host of confounders all at once, according to the inferentialist, no feature of experimental design can guarantee that all potential errors, including those having to do with the measurement of outcomes, with data analysis and reporting, with publication and many other aspects of the inference, have been ruled out. Whether or not they have been ruled out remains a judgment that can only be made after the entire body of evidence has been consulted. To give an example, for some time there was a discrepancy between randomized and observational studies in the effect of hormone replacement therapy on breast cancer, with the randomized studies showing a smaller risk than the observational studies. The reason was not, however, that the observational studies were inherently less reliable whereas the randomized studies got it right. The difference lay instead in the timing of the studies: the women in the randomized studies had on average been longer in menopause before starting the treatment. Reanalyzing the data from the randomized trials by adjusting for this temporal gap, the results fell in line and confirmed those of the observational studies (Vandenbroucke 2009). This outcome has nothing to do with the design of either trial or observational study and could only be reached by a systematic review and analysis of all the evidence.

An important issue that has been widely discussed among philosophers of the biomedical sciences concerns the role that evidence about mechanisms plays in causal inference. Risk factors do not cause medical outcomes across spatio-temporal gaps but through continuous biological pathways or "mechanisms" (sometimes also called "modes of action"). There is no doubt that understanding these mechanisms greatly enhances biomedical knowledge and is useful for numerous purposes, including the explanation of medical outcomes, improving intervention strategies, more accurate prognosis, and many more. According to one view, evidence about mechanisms is an important ingredient in successful causal inference (Russo and Williamson 2007). One reason to maintain this view is that causal conclusions can be regarded as always underdetermined by evidence about population-level associations because confounders cannot conclusively be ruled out. As confounders cannot always reliably be measured, and it is always possible that there are unanticipated confounders, causal conclusions should not rest on evidence about correlations alone.

Prima facie, the view discussed in the last paragraph is opposed to experimentalism but can be supported by inferentialism. Experimentalists maintain that confounders are ruled out by the design of a randomized trial. Accordingly, evidence about mechanisms plays only a small role if any in the so-called hierarchies of evidence used in evidence-based medicine. Inferentialists, by contrast, make causal judgments on the basis of evidence from a variety of sources, and evidence about mechanisms naturally fits into a diverse body of evidence. (For a discussion of the diverse roles that evidence about mechanisms can play in causal inference, see Clarke et al. 2014.)

Things are not quite so simple, however. On the one hand, it can be argued that knowledge about mechanisms is necessary in the planning and design of a randomized trial as well as in the analysis and interpretation of data (La Caze 2011). Accordingly, even if the immediate basis for a causal inference is an association generated by the experiment, the inference would not be reliable unless made against a backdrop of knowledge about mechanisms. On

the other hand, inferentialists need not require that this kind of knowledge be part of their diverse body of evidence. They can argue that while evidence about mechanisms can play a role in eliminating alternative hypotheses, there is no guarantee that it is a necessary ingredient (Reiss 2012).

Evidence about mechanisms is closely related to Hill's criterion (6), "plausibility." About it he says, "this is a feature I am convinced we cannot demand. What is biologically plausible depends upon the biological knowledge of the day" (Hill 1965). It is a simple fact about the history of medicine that our knowledge about biological pathways has increased dramatically since Hill wrote this. Examining the *IARC Monographs on the Evaluation of Carcinogenic Risk* over time, for instance, we find that in the 1970s there are practically no mechanistic considerations (e.g., IARC 1974b), whereas in the more recent edition the volume of reported data on mechanisms is larger than either that of epidemiological data or of data on animal models (e.g., IARC 2012). Concomitantly, we find an increasing number of substances that have been reclassified mainly on the basis of evidence about mechanisms. At the same time, at least for now it seems false to say that evidence about mechanisms is necessary for reliable causal inference. Many substances have been classified as carcinogenic for decades, and the original judgments were made without the benefit of data on mechanisms. It is not frequently the case that such judgments are overturned once the mechanistic data is in.

3. Interpreting Causal Claims

Once a causal claim has been established, what has been learned? What do we mean when we say that some risk factor causes a medical outcome? Philosophers distinguish five broad families of theories of causation: regularity, probabilistic, counterfactual, interventionist, and mechanistic. These theories provide interpretations of causal claims because they define the term "cause" that occurs in them. All five theories play, or have played, important roles in medicine.

Under the regularity view, a factor *causes* an outcome if and only if it is a necessary condition, a sufficient condition, both a necessary and sufficient condition, or an insufficient but non-redundant part of an unnecessary but sufficient (INUS) condition for the outcome. According to K. Codell Carter, early 19th-century medicine took a leap forward by adopting the regularity view in the guise of an "etiological viewpoint," the belief that

> diseases are best controlled and understood by means of causes, and in particular, by causes that are *natural* (that is, they depend on forces of nature as opposed to the willful transgression of moral or social norms), *universal* (that is, the same cause is common to every instance of a given disease), and *necessary* (that is, a disease does not occur in the absence of its cause).
>
> (Carter 2003: 1; emphasis in original)

The causes of large numbers of bacterial, viral, and deficiency diseases can be understood this way. Koch's postulates, which originate in the late 19th century, still embody the etiological viewpoint:

(i) The microorganism must be found in abundance in all organisms suffering from the disease, but should not be found in healthy organisms.
(ii) The microorganism must be isolated from a diseased organism and grown in pure culture.
(iii) The cultured microorganism should cause disease when introduced into a healthy organism.

(iv) The microorganism must be reisolated from the inoculated, diseased experimental host and identified as being identical to the original specific causative agent.

Postulate (iii) makes the presence of microorganism also sufficient for the disease. This basic understanding of the regularity view as a sufficient understanding of medical causation came under pressure as medical progress and an emerging interest in chronic (as opposed to infectious) diseases around the beginning of the 20th century challenged the understanding of cause as a necessary and/or sufficient condition. Koch knew that postulate (iii) is too strong (when read as a sufficient condition), because he found asymptomatic carriers of cholera (Koch 1893). Moreover, most causes of chronic diseases such as cancers or cardiovascular diseases are neither necessary nor sufficient conditions. The IARC classifies vinyl chloride as a human carcinogen, but exposure to the compound does not always lead to angiosarcomas (or other cancers), nor are all angiosarcomas (much less all cancers) caused by it.

The contemporary version of the regularity view maintains that causes are INUS conditions, which deals with some of the earlier problems. In contemporary epidemiology, for instance, causes are sometimes represented by the use of pie charts in which every wedge represents a condition that is necessary in the circumstances for disease; the entire pie represents the complex of conditions that are jointly sufficient (Rothman 1976). According to the INUS view, causes are neither necessary nor sufficient for their effects. Causes produce their outcomes only in conjunction with additional factors. Smoking, say, on its own is not followed by cancer. At minimum, a smoker has to be genetically susceptible and live long enough for the cancer to develop. Nor are most causal factors found in every instance of the disease as there are numerous alternative causes (each of which requires additional factors to produce the effect).

Both the original etiological viewpoint and the contemporary understanding of causes as INUS conditions are wedded to determinism: they assume that if all the conditions for an outcome are in place, the outcome will happen. Developments in the foundations of physics that occurred in the early 20th century led many researchers to abandon universal determinism and influenced thinking about causality in the biomedical sciences. Causes were no longer understood as sufficient for their effects (singly or jointly with additional factors) but rather as affecting merely the chances of outcomes, and thus the regularity view is no longer adequate.

According to the probabilistic view of causality, a cause is a factor that raises the probability of its effect in a causally homogenous population. The latter qualification is needed to distinguish between direct and confounded causal relations. The latter are cases where both the apparent cause and effect are in fact independent effects of a common cause. For example, under Fisher's constitutional hypothesis, smoking is not a cause of lung cancer but rather a byproduct of a common genetic factor. To distinguish between the two cases, we divide the population into two groups, one in which the genetic factor is present and one in which it is absent. If smoking raises the probability of lung cancer in both groups (assuming that the genetic factor is the only potential common cause), then it is a causal factor. A population is thus said to be causally homogenous whenever there is no variation among the causes of an outcome of interest in the population. If sex is a causally relevant factor, a causally homogenous population is one in which every member is a woman or one in which every member is a man; if age is a causally relevant factor, a causally homogenous population is one in which every member is in the same age group and so on.

It is important to note that the adequacy of the probabilistic view does not depend on whether factors such as Fisher's genetic condition are known or measurable. The question

is whether a factor in fact raises the probability in a causally homogeneous population, not whether there are means to test this. A conceptual rather than practical or epistemic question is, however, whether to demand that a cause raise the probability in *all* causally homogenous populations, in *some* populations, or *on average*. Nancy Cartwright has defended a requirement of "contextual unanimity" according to which only those factors that raise the probability in *all* causally homogenous populations are causes (Cartwright 1979); Brian Skyrms has a slightly weaker requirement according to which the causal factors raise the probability of the effect in *some* populations but do not lower it in any (Skyrms 1980); and John Dupré has argued that factors that raise the probability of their effects *on average* should be called causes (Dupré 1984). Contextual unanimity is a very strong requirement. If, say, there is a gene that makes some people immune to vinyl chloride, then the substance is not to be regarded as carcinogenic even if it raises the probability of cancer for most people. This point, and a look to biomedical practice, led Dupré to abandon it for a focus on average probability increases. What randomized trials establish, according to Dupré, are average causal effects, not contextually unanimous causes.

On the other hand, a disadvantage of calling a factor that raises the probability of their effects only on average a cause is that the status of a factor as cause depends on the actual distribution of factors in a population. Suppose that although VC exposure increases the risk of developing angiosarcoma in most people, a cancer immunization gene *lowered* the probability that those people exposed to VC who have the gene will develop angiosarcomas below that of the general population. That is, for these people, exposure actually protects them against cancer. Dupré would call VC carcinogenic only as long as relatively few people in the population had that gene; if more people had it, the substance would cease to be a cause or even become a preventer of cancer, even for those people without the protective gene. If one believes that whether or not VC is carcinogenic has to do with its intrinsic properties and the intrinsic properties of the person exposed, then this is an unwelcome result. Dupré's interpretation of causal claims is also prone to yield bad advice. A treatment that is effective on average may be harmful to some. If so, to learn that the treatment causes relief is misleading for those subpopulations whose members are harmed by it.

A third view of causality starts from individuals rather than populations. It maintains essentially that the treatment or risk factor is a cause of the medical outcome whenever the outcome would not have occurred if it had not been for the treatment or risk factor. For example, exposure to VC is the cause of a worker's angiosarcoma because he would not have developed the disease had he not been exposed to the substance. In philosophy, this counterfactual theory of causality, which originates in David Lewis' (1973) paper, has received much attention. Lewis and his followers have never found an empirical measure to determine the truth value of such counterfactuals, however. This job has been left to biostatisticians, who have developed the so-called potential-outcomes framework of causality. In that framework, the most fundamental quantity is the individual causal effect (ICE), which is defined as:ICE:

$$Y_t(u) - Y_c(u),$$

where Y measures the medical outcome of u (the patient), and t and c refer to treatment and control status, respectively. Thus, the individual causal effect measures the difference between the value the outcome variable would have assumed had the subject been treated and the value the variable would have assumed had the (same) subject not been treated. In theory, both values are counterfactual in nature. In practice, a patient can either be treated or not, but not both treated and not treated. Therefore, for any given subject u, we can

observe only either $Y_t(u)$ or $Y_c(u)$. Much of the literature on this framework develops strategies for identifying the individual causal effect or related quantities from observable data (e.g., Imbens and Rubin 2015).

The fourth, interventionist, account of causation was characterized in the previous section, when we discussed Woodward. Despite its enormous influence in philosophy, it has not been applied much in the biomedical sciences except in the area of psychology (see, for instance, the papers in Part I of Gopnik and Schulz 2007).

The final view of causality I will discuss is the mechanistic account, according to which a risk factor causes a medical outcome if, and only if, it produces the outcome by a mechanism of the appropriate kind. Many definitions of a mechanism have been advanced; according to one (Machamer et al. 2000: 3): "mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up conditions to finish or termination conditions." "Being productive of changes" is itself a causal notion, which is why some commentators have questioned the usefulness of this kind of account for understanding causality (e.g., Parascandola and Weed 2001). But arguably, understanding is provided by accounts of the specific entities of which mechanisms are composed and the activities in which they are engaged (Machamer 2004). At any rate, according to the mechanistic view, causal relations do not necessarily all share characteristics such as regular co-occurrence, probability raising, or counterfactual dependence, which is why proponents of the mechanistic view refuse to define causality in these terms.

Given this multiplicity of accounts of causation, a researcher who seeks guidance in interpreting causal claims is spoilt for choice, so much is clear. What is dubious, to say the least, is whether this is a good state of affairs. After all, confusion and worse can arise when different researchers interpret causal claims differently, especially when these claims are used for further research, treatment decisions, and health policy making. For example, if a team of researchers establishes that VC is a cause of angiosarcomas in the probabilistic sense, but the claim is mistaken by a jury or judge to be a counterfactual causal claim, then a firm that exposed workers to VC, some of whom subsequently developed the disease, may be held responsible because the jury or judge thinks that without having been exposed to the substance, they would not have gotten ill. The counterfactual claim, however, does not follow from the probabilistic causal claim.

There are at least three strategies to deal with this plurality of causal interpretations. The first is to try to reduce one conception to another. If, say, it can be proved that everything that is true about causality under a probabilistic view is equally true under an interventionist view and vice versa, then we can safely ignore one of these views. Unfortunately, all attempts to do so have proved unsuccessful. (See, e.g., Hausman and Woodward 1999; Cartwright 2006.) The second strategy is to put one's foot down and maintain that one's favorite account exhausts the meaning of "cause." The main problem with this strategy is that there are counterexamples to each account; that is, there are bona fide cases of causation that do not count as such under the given account and cases the account regards as causal that are not accepted as causal by common intuitions or scientific practice. Counterexamples to the regularity account were discussed above; similar cases can be found for each of the other accounts (for a detailed review, see Reiss 2015):

- (Probabilistic) Not all causes raise the probability of their effects because a cause can be connected to its effect through more than one mechanism in such a way that positive and negative influences cancel on balance. Birth control pills are a cause of deep vein thrombosis (DVT) but also prevent pregnancy, itself a cause of DVT. On average, the two routes (positive contribution and prevention) might exactly cancel so that women on the pill have the same chance as those who are not to develop DVT.

- (Counterfactual) When two or more causes compete to bring about an effect, there may be causation without counterfactual dependence. A smoker who was also exposed to asbestos may develop lung cancer due to his smoking. However, had he not smoked, he might have developed the disease anyway, because of asbestos exposure.
- (Interventionist) Not all causal relations are invariant to interventions. Interventions may sometimes chance the causal structure on the basis of which a higher-level relationship holds. Antibiotics are effective in the treatment of bacterial infections. Used too often, populations can become resistant. In this case, the intervention destroys the causal relation that was aimed to be used to bring about an effect.
- (Mechanistic) Some outcomes are caused by absences. Vitamin-D deficiency can cause multiple sclerosis. However, absences are not mechanistically connected to their effects. (This does not mean that mechanisms are not used in the *explanation* of the onset of a disease caused by an absence, but the absence cannot be the starting point of a causal process that terminates in the effect.)

The third strategy is to maintain that causality is not a feature of the external world but rather one of reasoning agents such as scientific researchers. Information about regularities, probabilistic dependencies, results of experiments, and the like provides reasons to believe causal relationships, but causal claims do not (or need not) represent anything specific in the world. According to one such "subjective" theory of causality, epistemic causality, "[c]ausal relationships are to be identified with the causal beliefs of an omniscient rational agent" (Russo and Williamson 2007: 168). Thus, causality depends on what an agent believes in the ideal case in which he/she is in the possession of all relevant evidence (though it is acknowledged that real agents are rarely in that situation). Another subjective theory, the inferentialist theory (which is closely related to inferentialism about evidence that was discussed above):

> maintains that the meaning of causal claims is given by their inferential connections with other claims. In particular, causal claims are inferentially related to *evidential claims*—the claims from which a causal claim can be inferred—as well as to claims about future events, explanatory claims, claims attributing responsibility, and counterfactual claims (claims predicting "what would happen if")—the claims that can be inferred from a causal claim.
>
> (Reiss 2015a: 20; emphasis in original)

Subjective theories do not suffer from the drawbacks of the other two strategies. They neither try to reduce one feature of causal relationships to another, nor do they suffer from obvious counterexamples (see, for instance, Reiss 2015a, Chapter 5, about how the inferentialist theory deals with cases of redundant causation which pose problems for many standard theories). They also address the original challenge adequately. The epistemic theory holds that causal claims have, in principle, one unique meaning, namely, what an omniscient rational agent believes to be true. The inferentialist theory can be understood as giving up looking for a definition of cause, instead allowing that there may be multiple notions, and asking about what kind of practices provide evidence for causal claims, what inferences from this evidence are justified, and what purposes knowledge of causal claims in medicine serves. It is a pragmatic theory that starts with a medical or policy problem, asks what kinds of causal knowledge are relevant for addressing the problem and what kinds of evidence are needed to substantiate this knowledge without assuming that meaning necessarily carries over from one context to the next.

4. On the Usefulness of Causality in Medicine

In the introductory section, I suggested that biomedical researchers seek knowledge of causal claims because it is useful in attaining the discipline's more ultimate purposes such as explanation, prediction, and making successful treatment and health policy decisions. We may ask, however, whether knowledge of causal claims really does promote these more ultimate purposes. In particular, we may ask whether knowledge of causal claims is necessary or sufficient (or both) for explanation, prediction, and decision making.

It turns out that the answer is ambiguous because it depends on context and on the kind of causal claim that is being used in that context. It is certainly true that causal claims can help explain outcomes, but their ability to do so depends on the precise nature of the explanatory interest and the causal claim at hand. "Vinyl chloride causes cancer of the liver" may be cited to explain a particular clustering of incidences (along with information about exposure), but it hardly explains why this worker rather than that one developed the disease, or why many workers who were exposed do not develop it or some non-exposed people do. Without information about the biological pathways through which they produce neoplasms, causal explanations are very thin at best.

The relation between causality and prediction is even more tenuous. When relationships are stable, we can successfully predict on the basis of correlations—knowledge of the true causal structure is not necessary. On the other hand, when relations are not stable (for instance, because the composition of causal mechanisms changes, some mechanisms operate indeterministically, or interferences occur), knowledge of causal relations does not help much to improve predictability. If, say, exposure to VC produces a stable number of liver cancers in a population and gives all exposed workers a specific set of symptoms, such as peripheral neuropathy and pain in the fingers, we can use information about the symptoms (a correlate) to make a prediction about the chances of developing liver cancer. If, on the other hand, the relationships are not stable, making a prediction on the basis of observing the real cause is prone to be unreliable. It is certainly possible that the relationship between some indicator, which is not directly causally connected to an outcome, and the outcome is more stable than the relationship between a cause and its effect. The point is that for prediction, stability is important, not causality.

Lastly, both of these problems occur in the case of decision making. For good decisions, we first need the right kind of causal knowledge. If some treatment does not operate in a "contextually unanimous" fashion (see previous section), then knowing that it is effective at the population level will not be a good basis for an individual treatment recommendation, even if that individual is a member of that population. Similarly, as Alex Broadbent has pointed out, that *C* (e.g., some risk factor) causes *E* (e.g., an adverse medical outcome) does not mean that reducing *C* *will* reduce *E* (Broadbent 2013). A ban on sugary drinks will not necessarily lead to lower obesity or diabetes rates because it depends on what people do instead. If, as does not seem implausible, people substitute an equally or more risky behavior, rates may stay put or increase even though the policy was based on a genuine causal relationship. Once again, in the policy context we want stable relationships between a policy variable and an outcome, and what causal relationships, if any, these are based on is immaterial.

Broadbent argues that epidemiologists are primarily interested in explanation and prediction, not in causation (Broadbent 2013). He bemoans that philosophers of science have basically ignored prediction as a topic of methodological analysis. The above-mentioned considerations support this view.

This is not to argue, of course, that knowledge of causal claims is not useful. An important take-home lesson, however, is that biomedical researchers are seldom interested in causal

claims in their own right but rather because, and to the extent to which, they help to attain more ultimate purposes. Whether they do so or not is a contextual matter that has to do with the precise nature of the purpose pursued as well as the causal claim in question.

References

Broadbent, Alex (2013). *The Philosophy of Epidemiology*. London: Palgrave Macmillan.

Carter, K. Codell (2003). *The Rise of Causal Concepts of Disease: Case Histories*. Aldershot: Ashgate.

Cartwright, Nancy (1979). "Causal Laws and Effective Strategies." *Noûs* 13:419–437.

——— (2006). "From Metaphysics to Method: Comments on Manipulability and the Causal Markov Condition." *British Journal for Philosophy of Science* 57(1):197–218.

Clarke, Brendan, Donald Gillies, Phyllis Illari, Federica Russo and Jon Williamson (2014). "Mechanisms and the Evidence Hierarchy." *Topoi* 33:339–360.

Dupré, John (1984). "Probabilistic Causality Emancipated." In *Midwest Studies in Philosophy IX*, ed. P. French, T. Uehling Jr and H. Wettstein. 169–175. Minneapolis (MS): University of Minnesota Press.

Gopnik, Alison and Laura Schulz (2007). *Causal Learning: Psychology, Philosophy, and Computation*. New York: Oxford University Press.

Hausman, Daniel and James Woodward (1999). "Independence, Invariance and the Causal Markov Condition." *British Journal for the Philosophy of Science* 50:521–583.

Hill, Austin Bradford (1965). "The Environment and Disease: Association or Causation?" *Proceedings of the Royal Society of Medicine* 58(5):295–300.

IARC (1974a). *IARC Monographs on the Evaluation of Carcinogenic Risk of Chemicals to Man*. Volume 7: Some anti-thyroid and related substances, nitrofurans and industrial chemicals. Lyon: International Agency for Research on Cancer.

——— (1974b). *IARC Monographs on the Evaluation of Carcinogenic Risk of Chemicals to Man*. Volume 14: Asbestos. Lyon: International Agency for Research on Cancer.

——— (2012). *IARC Monographs on the Evaluation of Carcinogenic Risk to Humans*. Volume 106: Trichloroethylene, tetrachloroethylene, and some other chlorinated agents. Lyon: International Agency for Research on Cancer.

——— (2014). *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*. Volume 106: Trichloroethylene, tetrachloroethylene, and some other chlorinated agents. Lyon: International Agency for Research on Cancer.

Imbens, Guido W. and Donald Rubin (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. New York (NY): Cambridge University Press.

Koch, Robert (1893). "Über den augenblicklichen Stand der bakteriologischen Choleradiagnose." *Zeitschrift für Hygiene und Infectionskrankheiten* 14:319–333.

La Caze, Adam (2011). "The Role of Basic Science in Evidence-Based Medicine." *Biology and Philosophy* 26(1):81–98.

Lewis, David (1973). *Counterfactuals*. Cambridge (MA): Harvard University Press.

Machamer, Peter (2004). "Activities and Causation: The Metaphysics and Epistemology of Mechanisms." *International Studies in the Philosophy of Science* 18(1):27–39.

Machamer, Peter, Lindley Darden and Carl Craver (2000). "Thinking About Mechanisms." *Philosophy of Science* 67:1–25.

Maltoni, C. and G. Lefemine (1974). "Carcinogenicity Bioassays on Vinyl Chloride: I. Research Plan and Early Results." *Environmental Research* 7:387–405.

Parascandola, Mark (2004). "Two Approaches to Etiology: The Debate Over Smoking and Lung Cancer in the 1950s." *Endeavour* 28(2):81–86.

Parascandola, M., & Weed, D. L. (2001). Causation in epidemiology. *Journal of Epidemiology and Community Health*, 55(12), 905–912.

Reiss, Julian (2012). "Third Time's a Charm: Wittgensteinian Pluralisms and Causation." In *Causality in the Sciences*, ed. P. McKay Illari, F. Russo and J. Williamson. 907–927. Oxford: Oxford University Press.

——— (2015). *Causation, Evidence, and Inference*. New York (NY): Routledge.

Rothman, Kenneth (1976). "Causes." *American Journal of Epidemiology* 104(6):587–592.
Russo, Federica and Jon Williamson (2007). "Interpreting Causality in the Health Sciences." *International Studies in the Philosophy of Science* 21(2):157–170.
Skyrms, Brian (1980). *Causal Necessity: Pragmatic Investigation of the Necessity of Laws*. New Haven (CT): Yale University Press.
Smith, Gordon and Jill Pell (2003). "Parachute Use to Prevent Death and Major Trauma Related to Gravitational Challenge: Systematic Review of Randomised Controlled Trials." *British Medical Journal* 327:1459–1461.
Vandenbroucke, Jan P. (2009). "The HRT Controversy: Observational Studies and RCTs Fall in Line." *Lancet* 373:1233–1235.
Woodward, James (2003). *Making Things Happen*. Oxford: Oxford University Press.

Further Reading

Agassi, Joseph (1976). "Causality and Medicine." *Journal of Medicine and Philosophy* 1(4):301–317.
Reiss, Julian (2015). "A Pragmatist Theory of Evidence." *Philosophy of Science* 82(3):341–362.
Worrall, John (2007). "Evidence in Medicine and Evidence-Based Medicine." *Philosophy Compass* 2:981–1022.